```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:     January 15, 2006, 12:59:30 ; Search time 171 Seconds
                                          (without alignments)
                                          800.423 Million cell updates/sec

Title:          US-10-751-235-1
Perfect score:  393
Sequence:       1 LQPYAEDGSAVNMEAKFSQM........STDLLPYWKIDALCKIVPRQ 77

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters:    2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10751235/runat_12012006_120916_2538/app_query.fasta_1.263
-DB=Issued_PatentS_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCU=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235 @CGN_1_1_193 @runat_12012006_120916_2538 -LONGLOG
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Issued_Patents_NA:*
         1:  /cgn2_6/ptodata/1/ina/1_COMB.seq:*
         2:  /cgn2_6/ptodata/1/ina/5_COMB.seq:*
         3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
         4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
         5:  /cgn2_6/ptodata/1/ina/H_COMB.seq:*
         6:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
         7:  /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
         8:  /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
         9:  /cgn2_6/ptodata/1/ina/backfile1.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

### SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 200 | 50.9 | 1831 | 3 | US-08-948-564-15 | Sequence 15, Appl |
| 2 | 196 | 49.9 | 404 | 3 | US-09-615-192A-156 | Sequence 156, App |
| 3 | 196 | 49.9 | 404 | 3 | US-09-169-789-156 | Sequence 156, App |
| 4 | 76 | 19.3 | 3150 | 3 | US-10-018-730A-3 | Sequence 3, Appl |
| 5 | 74 | 18.8 | 1633 | 3 | US-09-583-447A-7 | Sequence 7, Appli |
| 6 | 71.5 | 18.2 | 1349 | 3 | US-09-583-447A-5 | Sequence 5, Appli |
| 7 | 71.5 | 18.2 | 1515 | 3 | US-09-583-447A-3 | Sequence 3, Appli |
| 8 | 71.5 | 18.2 | 1659 | 3 | US-09-583-447A-1 | Sequence 1, Appli |
| 9 | 71.5 | 18.2 | 1973 | 3 | US-09-583-447A-11 | Sequence 11, Appl |
| 10 | 71 | 18.1 | 2352 | 2 | US-08-802-141-3 | Sequence 3, Appli |
| 11 | 71 | 18.1 | 6176 | 3 | US-08-911-321-6 | Sequence 6, Appli |
| 12 | 69.5 | 17.7 | 432 | 3 | US-09-248-796A-4077 | Sequence 4077, Ap |
| 13 | 69 | 17.6 | 2217 | 3 | US-09-712-363-92 | Sequence 92, Appl |
| 14 | 69 | 17.6 | 2363 | 2 | US-07-603-133B-19 | Sequence 19, Appl |
| 15 | 69 | 17.6 | 2368 | 3 | US-08-653-740-2 | Sequence 2, Appli |
| 16 | 69 | 17.6 | 2368 | 2 | US-09-073-594-2 | Sequence 2, Appli |
| 17 | 69 | 17.6 | 2368 | 3 | US-09-275-925-2 | Sequence 2, Appli |
| c 18 | 69 | 17.6 | 4403765 | 3 | US-09-103-840A-2 | Sequence 2, Appli |
| c 19 | 69 | 17.6 | 4411529 | 3 | US-09-103-840A-1 | Sequence 1, Appli |
| 20 | 65.5 | 16.7 | 2338 | 3 | US-09-336-115C-21 | Sequence 21, Appl |
| 21 | 65.5 | 16.7 | 2429 | 3 | US-09-336-115C-11 | Sequence 11, Appl |
| 22 | 65.5 | 16.7 | 73788 | 3 | US-09-949-016-12358 | Sequence 12358, A |
| 23 | 65 | 16.5 | 629 | 2 | US-08-454-557C-43 | Sequence 43, Appl |
| 24 | 65 | 16.5 | 629 | 3 | US-08-340-426D-43 | Sequence 43, Appl |
| 25 | 65 | 16.5 | 629 | 2 | US-08-450-673C-43 | Sequence 43, Appl |
| 26 | 65 | 16.5 | 629 | 6 | PCT-US95-17111A-43 | Sequence 43, Appl |
| 27 | 65 | 16.5 | 1431 | 3 | US-09-268-992-6 | Sequence 6, Appli |
| 28 | 65 | 16.5 | 1431 | 3 | US-09-657-474-6 | Sequence 6, Appli |
| 29 | 65 | 16.5 | 1485 | 3 | US-09-268-992-5 | Sequence 5, Appli |
| 30 | 65 | 16.5 | 1485 | 3 | US-09-657-474-5 | Sequence 5, Appli |
| 31 | 65 | 16.5 | 1957 | 3 | US-09-268-992-3 | Sequence 3, Appli |
| 32 | 65 | 16.5 | 1957 | 3 | US-09-657-474-3 | Sequence 3, Appli |
| 33 | 65 | 16.5 | 2055 | 3 | US-09-268-992-1 | Sequence 1, Appli |
| 34 | 65 | 16.5 | 2055 | 3 | US-09-657-474-1 | Sequence 1, Appli |
| 35 | 64.5 | 16.4 | 601 | 3 | US-09-949-016-79444 | Sequence 79444, A |
| 36 | 64.5 | 16.4 | 784019 | 3 | US-09-949-016-14033 | Sequence 14033, A |
| 37 | 64.5 | 16.4 | 828152 | 3 | US-09-949-016-12777 | Sequence 12777, A |
| 38 | 64 | 16.3 | 804 | 3 | US-09-710-279-1611 | Sequence 1611, Ap |
| 39 | 64 | 16.3 | 861 | 3 | US-09-134-001C-2487 | Sequence 2487, Ap |
| 40 | 64 | 16.3 | 1707 | 3 | US-09-022-655-1060 | Sequence 1060, Ap |
| 41 | 64 | 16.3 | 1707 | 3 | US-09-949-016-121 | Sequence 121, App |
| c 42 | 64 | 16.3 | 1707 | 3 | US-09-949-016-1220 | Sequence 1220, Ap |
| 43 | 64 | 16.3 | 2918 | 3 | US-09-710-279-3932 | Sequence 3932, Ap |
| 44 | 64 | 16.3 | 3660 | 3 | US-09-710-279-4116 | Sequence 4116, Ap |
| 45 | 63.5 | 16.2 | 1158 | 3 | US-09-543-681A-3780 | Sequence 3780, Ap |

### ALIGNMENTS

```
RESULT 1
US-08-948-564-15
; Sequence 15, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
;    APPLICANT: Siminszky, Balazs
;    APPLICANT: Dewey, Ralph E.
;    APPLICANT: Corbin, Frederick T.
;    TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
;    TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
;    NUMBER OF SEQUENCES: 23
;    CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Virginia C. Bennett
;     STREET: PO Box 37428
;     CITY: Raleigh
;     STATE: No. 6121512th Carolina
;     COUNTRY: USA
;     ZIP: 27627
;    COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/948,564
;     FILING DATE:
;     CLASSIFICATION: 800
;    ATTORNEY/AGENT INFORMATION:
;     NAME: Bennett, Virginia C.
;     REGISTRATION NUMBER: 37,092
;     REFERENCE/DOCKET NUMBER: 5051-409
;    TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 919-854-1400
```

```
;   TELEFAX:  919-854-1401
; INFORMATION FOR SEQ ID NO:  15:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  1831 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  20..1747
US-08-948-564-15

Alignment Scores:
Pred. No.:              2.57e-20       Length:            1831
Score:                  200.00         Matches:           40
Percent Similarity:     73.24%         Conservative:      12
Best Local Similarity:  56.34%         Mismatches:        19
Query Match:            50.89%         Indels:            0
DB:                     3              Gaps:              0

US-10-751-235-1 (1-77) x US-08-948-564-15 (1-1831)

Qy    7 AspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGly  26
          ||    :::::::::::||||||||||| ::: :||||||||:::||||||
Db  647 GACTCAATTGAATTGGATCTGGAGGCAGAGTTTTCTAGTTGGCTCTTGATATTATTGGG  706

Qy   27 LeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAla  46
         |||  :::||||||:||||||||::::|||            :::||||||||||||:::
Db  707 CTTGGTGTTGTTCAACTATGACTTTGGTTCTGTCACCAAAGAATCTCCAGTTATTAAGGCA  766

Qy   47 ValTyrThrAlaLeuLysGluAlaGlyLeuArgSerThrAspLeuLeuProTyrTrpLys  66
         ||||| |||   ||||||||||| |||||| ||||||||||:::||||||||||||||||
Db  767 GTCTATGGCACTCTTTTGAAGGTGAACACAGATCCACTTTCTACATTCCATATTGGAAA  826

Qy   67 IleAspAlaLeuCysLysIleValProArgGln  77
          ||||||||||||||||||||||||||||||
Db  827 ATTCCATTGGCAAGGTGGATAGTCCCAAGGCAA  859

RESULT 2
US-09-615-192A-156
; Sequence 156, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
;   APPLICANT:  Bloksberg, Leonard N.
;   APPLICANT:  Havukkala, Ilkka
;   TITLE OF INVENTION:  Materials and Methods for the
;   TITLE OF INVENTION:  Modification of Plant Lignin Content
;   FILE REFERENCE:  11000.1003c4U
;   CURRENT APPLICATION NUMBER: US/09/615,192A
;   CURRENT FILING DATE:  2000-07-12
;   PRIOR APPLICATION NUMBER: US 08/975,316
;   PRIOR FILING DATE:  1997-11-21
;   PRIOR APPLICATION NUMBER: US 08/713,000
;   PRIOR FILING DATE:  1996-09-11
;   PRIOR APPLICATION NUMBER: US 09/169,789
;   PRIOR FILING DATE:  1998-10-09
;   NUMBER OF SEQ ID NOS:  405
;   SOFTWARE:  FastSEQ for Windows Version 3.0
;   SEQ ID NO 156
;     LENGTH:  404
;     TYPE:  DNA
;     ORGANISM:  Pinus radiata
US-09-615-192A-156

Alignment Scores:
Pred. No.:              1.26e-20       Length:            404
Score:                  196.00         Matches:           39
Percent Similarity:     74.63%         Conservative:      11
Best Local Similarity:  58.21%         Mismatches:        17
Query Match:            49.87%         Indels:            0
DB:                     3              Gaps:              0
```

```
US-10-751-235-1 (1-77) x US-09-615-192A-156  (1-404)

Qy   11 ValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeuSerLeuPhe  30
          ||||||||||||||||| ::||    ||||||| :::|||||||
Db   31 GTTGACATGGAGGCAGAATTTCAAATCTAGCTTGGACATTATTGGATTGTGTGTATTT  90

Qy   31 AsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaValTyrThrAla  50
         ||||||:::||  ||:::||| :::::||||  ||||||||||:::|||||||:|||||
Db   91 AACTATGATTTGGATCCGTTACTCGAGAATCACCAGTAATCAAGGCAGTCTATGGTACA  150

Qy   51 LeuLysGluAlaGlyLeuArgSerThrAspLeuLeuProTyrTrpLysIleAspAlaLeu  70
          ||||| ::: ||||||| :::|||||||||||||||||||
Db  151 TTGTTTGAAGCTGAGCATAGATCAACCTTTTACATACCATACTGGAAATTCCGGCTGGCA  210

Qy   71 CysLysIleValProArgGln  77
          :::||||||||||||||
Db  211 AGATGGTTAGTTCCTCGCCAA  231

RESULT 3
US-09-169-789-156
; Sequence 156, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
;   APPLICANT:  Bloksberg, Leonard N.
;   APPLICANT:  Havukkala, Ilkka
;   TITLE OF INVENTION:  Materials and Methods for the
;   TITLE OF INVENTION:  Modification of Plant Lignin Content
;   FILE REFERENCE:  11000.1003c2
;   CURRENT APPLICATION NUMBER: US/09/169,789
;   CURRENT FILING DATE:  1998-10-09
;   EARLIER APPLICATION NUMBER: US 08/975,316
;   EARLIER FILING DATE:  1997-11-21
;   EARLIER APPLICATION NUMBER: US 08/713,000
;   EARLIER FILING DATE:  1996-09-11
;   NUMBER OF SEQ ID NOS: 185
;   SOFTWARE:  FastSEQ for Windows Version 3.0
;   SEQ ID NO 156
;     LENGTH: 404
;     TYPE: DNA
;     ORGANISM: Pinus radiata
US-09-169-789-156

Alignment Scores:
Pred. No.:              1.26e-20       Length:            404
Score:                  196.00         Matches:           39
Percent Similarity:     74.63%         Conservative:      11
Best Local Similarity:  58.21%         Mismatches:        17
Query Match:            49.87%         Indels:            0
DB:                     3              Gaps:              0

US-10-751-235-1 (1-77) x US-09-169-789-156  (1-404)

Qy   11 ValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeuSerLeuPhe  30
          ||||||||||||||||| ::||    ||||||| :::|||||||
Db   31 GTTGACATGGAGGCAGAATTTCAAATCTAGCTTGGACATTATTGGATTGTGTGTATTT  90

Qy   31 AsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaValTyrThrAla  50
         ||||||:::||  ||:::||| :::::||||  ||||||||||:::|||||||:|||||
Db   91 AACTATGATTTGGATCCGTTACTCGAGAATCACCAGTAATCAAGGCAGTCTATGGTACA  150

Qy   51 LeuLysGluAlaGlyLeuArgSerThrAspLeuLeuProTyrTrpLysIleAspAlaLeu  70
          ||||| ::: ||||||| :::|||||||||||||||||||
Db  151 TTGTTTGAAGCTGAGCATAGATCAACCTTTTACATACCATACTGGAAATTCCGGCTGGCA  210

Qy   71 CysLysIleValProArgGln  77
          :::||||||||||||||
Db  211 AGATGGTTAGTTCCTCGCCAA  231

RESULT 4
US-10-018-730A-3
; Sequence 3, Application US/10018730A
; Patent No..6794168
; GENERAL INFORMATION:
```

*SUBSET OF STIC SEQUENCE SEARCH*
*for SEQ ID NO. 5*

```
KW   seed protein yield.
OS   Viridiplantae.
XX
PN   US2004216190-A1.
XX
PD   28-OCT-2004.
XX
PF   18-DEC-2003; 2003US-00739930.
XX
PR   28-APR-2003; 2003US-00424599.
PR   28-APR-2003; 2003US-00425115.
XX
PA   (KOVA/) KOVALIC D K.
XX
PI   Kovalic DK;
XX
DR   WPI; 2004-757369/74.
XX
PT   New recombinant DNA constructs useful in the field of biochemistry and
PT   genetics, and in particular for producing transgenic plants with improved
PT   biological characteristics.
XX
PS   Claim 1; SEQ ID NO 773; 14pp; English.
XX
CC   The invention relates a recombinant DNA construct comprising a
CC   polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC   1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC   (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC   Arabidopsis, wheat and rape but the specification does not indicate which
CC   sequences is derived from which organism. Also included is a method of
CC   producing a plant having an improved property, comprising transforming a
CC   plant with a recombinant DNA construct comprising a promoter region
CC   functional in a plant cell operably joined to a polynucleotide encoding a
CC   polypeptide associated with the property, and growing the transformed
CC   plant. The property is selected from improving plant cold tolerance, for
CC   manipulating growth rate in plant cells by modification of the cell cycle
CC   pathway, for improving plant drought tolerance, for providing increased
CC   resistance to plant disease, for galactomannan production, for production
CC   of plant growth regulators, for improving plant heat tolerance, for
CC   improving plant tolerance to herbicides, for increasing the rate of
CC   homologous recombination in plants, for lignin production, for improving
CC   plant tolerance to extreme osmotic conditions, for improving plant
CC   tolerance to pathogens or pests, for yield improvement by modification of
CC   photosynthesis, for modifying seed oil yield and/or content, for
CC   modifying seed protein yield and/or content, for yield improvement by
CC   modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC   and for yield improvement by providing improved plant growth and
CC   development under at least one stress condition. The polynucleotide may
CC   also encode a plant transcription factor. The methods and compositions of
CC   the present invention are useful in the field of biochemistry and
CC   genetics, in particular for producing transgenic plants with improved
CC   biological characteristics such as increased yield, improved nitrogen
CC   flow, increasing plant tolerance to cold or heat, improving plant
CC   tolerance to extreme osmotic and drought conditions, and improving plant
CC   tolerance to plant pests or pathogens. They can also be used in physical
CC   arrays of molecules, plant breeding markers, computer-based storage and
CC   analysis systems. The present sequence is one of the 5544 plant cDNA
CC   sequences of the invention. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ   Sequence 3377 BP; 967 A; 729 C; 736 G; 945 T; 0 U; 0 Other;

Query Match         32.5%;    Score 803;   DB 13;   Length 3377;
Best Local Similarity   73.2%;   Pred. No. 1.1e-208;
Matches 1347;   Conservative 0;   Mismatches 0;   Indels 494;   Gaps 5;

Qy    1 ATGGAGTCTTCACTCTTTTCTCCATCTCCTCTTCTCTCTTCTTCTCTTACTCTGCAAAA   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   79 ATGGAGTCTTCACTCTTTTCTCCATCTCCTCTTCTCTCTTCTTCTCTTACTCTGCAAAA  138
```

```
Qy 1681 TCTCTAATTATTCATGAACTAAAATTTCTGATTGATTGTTTCCTGGTAGTCTTGATAA 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 TCTCTAATTATTCATGAACTAAAATTTCTGATTGATTGTTTCCTGGTAGTCTTGATAA 1740

Qy 1741 GAAGAGCTCAAGTTCCTGACATTCTTCCTGGAACTATAAGGTCAATACCGGACAAGACA 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 GAAGAGCTCAAGTTCCTGACATTCTTCCTGGGAACTATAAGGTCAATACCGGACAAGACA 1800

Qy 1801 TTATGATTTCAGTCTATAACATCCATCGTCTCCGAGGTACAGTTCTTCCTTCCTCTC 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 TTATGATTTCAGTCTATAACATCCATCGTCTTCCGAGGTACAGTTCTTCCTTCCTCTC 1860

Qy 1861 GTCCATAGTATAACATAGGGGAGCCTAATCCTTCTTCTTCCAATGATCTTTGTGTGGTTCGG 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 GTCCATAGTATAACATAGGGGAGCCTAATCCTTCTTCTTCAATGATCTTTGTGTGGTTCGG 1920

Qy 1921 ATATCTAACCGGAGTGGACATTCCTAGTATTACATTCATGCCCCACATTCTTATGTGTTT 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 ATATCTAACCGGAGTGGACATTCCTAGTATTACATTCATGCCCCACATTCTTATGTGTTT 1980

Qy 1981 GTTGTTTGTTATTCCAAAGGTATGGGAAAAAGCTGAGGAATTTCGCCTGAACGATTCGA 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 GTTGTTTGTTATTCCAAAGGTATGGGAAAAAGCTGAGGAATTTCGCCTGAACGATTCGA 2040

Qy 2041 CATAGAATGGCGCAATCCTAACGAAACAAACAAACTGATTCAAGTAAACTCAGTAGAACAC 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 CATAGAATGGCGCAATCCTAACGAAACAAACAAACTGATTCAAGTAAACTCAGTAGAACAC 2100

Qy 2101 ATCTTTGACAACAACTACTGAATCAAGATTAGTTGGTTTGATTAGGGAATTTAAAAGAT 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 ATCTTTGACAACAACTACTGAATCAAGATTAGTTGGTTTGATTAGGGAATTTAAAAGAT 2160

Qy 2161 GATTTCTTTTTCACCAGATTCATCCCATTCAGTCGAGGGCCTAGAAAATGTGTAGGCG 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 GATTTCTTTTTCACCAGATTCATCCCATTCAGTCGAGGGCCTAGAAAATGTGTAGGCG 2220

Qy 2221 ATCAGTTTGCATTGATTGGAGGCAATTGTGGCACTCGCGGTGTTCTTCAGCGGTTAAACG 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 ATCAGTTTGCATTGATTGGAGGCAATTGTGGCACTCGCGGTGTTCTTCAGCGGTTAAACG 2280

Qy 2281 TTGAGCTGGTTCCTGATCAGACCATTAGCATGTGCCACCAGGAGCAACCATACACACCACCA 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 TTGAGCTGGTTCCTGATCAGACCATTAGCATGCATGACCACCAGGAGCAACCATACACACCACCA 2340

Qy 2341 ATGTATGCCAATGTTCTCACACTCGAGAGATTAATGAGAGTGTCGTTTTGTTTTAGAATG 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 ATGTATGCCAATGTTCTCACACTCGAGAGATTAATGAGAGTGTCGTTTTGTTTTAGAATG 2400

Qy 2401 ATTCCAATTTCTCTAATGCTGATATTTTCAATTTCAGGGATTGTATATGAAGGTGAGCCA 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 ATTCCAATTTCTCTAATGCTGATATTTTCAATTTCAGGGATTGTATATGAAGGTGAGCCA 2460

Qy 2461 AAGGTAA 2467
         |||||||
Db 2461 AAGGTAA 2467
```

```
RESULT 4
ADT15447
ID   ADT15447 standard; cDNA; 3377 BP.
XX
AC   ADT15447;
XX
DT   13-JAN-2005  (first entry)
XX
DE   Plant cDNA, Seq ID 773.
XX
KW   Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW   drought tolerance; disease resistance; galactomannan production;
KW   plant growth regulator; heat tolerance; herbicide tolerance;
KW   lignin production; extreme osmotic condition tolerance;
KW   pathogens resistance; pest resistance; yield improvement; seed oil yield;
```

Qy    61  CCTACGGCGTCTTTATCACCAAAACCCAAATTCACATTCCTCCATCAGATCCTCCATTGAG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   139  CCTACGGCGTCTTTATCACCAAAACCCAAATTCACATTCCTCCATCAGATCCTCCATTGAG  198

Qy   121  AAACCCAAACCCAAACTCGAGACCAAACTCGAAATCCCAATCATGGGTCAGTCCCGAT  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   199  AAACCCAAACCCAAACTCGAGACCAAACTCGAAATCCCAATCATGGGTCAGTCCCGAT  258

Qy   181  TGGCTCACACACTCACTCGTACCCTTCCTCAGGAAAAAACGACGAGTCAGTATACCA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   259  TGGCTCACACACTCACTCGTACCCTTCCTCAGGAAAAAACGACGAGTCAGTATACCA  318

Qy   241  ATCGCGAACGCGAAGCTCGACGATGTCGGCTGATCTCCTCGGAGTGCTCTCTTCCTACCT  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   319  ATCGCGAACGCGAAGCTCGACGATGTCGGCTGATCTCCTCGGAGTGCTCTCTTCCTACCT  378

Qy   301  CTCTACAAATGGATGAAATGACTACGGACCCATTTACCGTCTCGCTGCTCGTGCTCCGTAAT  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   379  CTCTACAAATGGATGAAATGACTACGGACCCATTTACCGTCTCGCTGCTCGTGCTCCGTAAT  438

Qy   361  TTCGTAATTGTGAGCGACCCCAGCGATAGCTAAACATGTTTGAGGAATTATCCAAAGTAC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   439  TTCGTAATTGTGAGCGACCCCAGCGATAGCTAAACATGTTTGAGGAATTATCCAAAGTAC  498

Qy   421  GCTAAAGGCTTAGTCGCTGAAGTCTCTGAATTCTCATTTGGTTCGGGTTTCGCTATCGCT  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   499  GCTAAAGGCTTAGTCGCTGAAGTCTCTGAATTCTCATTTGGTTCGGGTTTCGCTATCGCT  558

Qy   481  GAAGGACCTCTTTGGACAGTAATTCATCTCCCTCTATCCAATTTTGAAGTTTTTGGAA  540
           ||||||||||||||||||||||||
Db   559  GAAGGACCTCTTTGGACA------------------------------------  576

Qy   541  TTGTGGAGTAATGTGACTGTCTTGTATGATAAGTAACTCTAATTTTAGGGTTTAGAT  600
Db   577  ------------------------------------------------------  576

Qy   601  TCCAATCTTCTCTATTGGGCTTAGCTGAAGTCTGATTTTTTACATAGGCGAGGCGTAGAG  660
                                                         ||||||||||||||
Db   577  --------------------------------------------GCGAGGCGTAGAG  589

Qy   661  CGGTGGTTCCATCGCTTCACAGGAGTATTGTCTGTGATTGTGGAGAGAGTATTCTGCA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   590  CGGTGGTTCCATCGCTTCACAGGAGTATTTGTCTGTGATTGTGGAGAGAGTATTCTGCA  649

Qy   721  AATGTGCAGAGAGGCTTGTTGAGAGATTGCAGCCTTATGCAGAAGACGGAAGTGCGTGA  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   650  AATGTGCAGAGAGGCTTGTTGAGAGATTGCAGCCTTATGCAGAAGACGGAAGTGCGTGA  709

Qy   781  ATATGGAAGCGCAAGTTCTCTCAGATGACACTTGATTGTCATTGGGTTGTCTCTTTTTAACT  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   710  ATATGGAAGCGCAAGTTCTCTCAGATGACACTTGATTGTCATTGGGTTGTCTCTTTTTAACT  769

Qy   841  ACAATTTCGATTCTTTGACTACTGATAGTCCGTGTCATTGAAGCTGTTTACACTGCTCTTA  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   770  ACAATTTCGATTCTTTGACTACTGATAGTCCGTGTCATTGAAGCTGTTTACACTGCTCTTA  829

Qy   901  AAGAAGCTGAGCTTCGTTCTCACTGATCTTCTGCCCATATTGGAAGGCAAGTTCCGTGTGTT  960
           ||||||||||||||||||||||
Db   830  AAGAAGCTGAGCTTCGTTCTCACTGATCTTCTGCCCATATTGGA-----------------  871

Qy   961  TTTTCTGTCGTGTTGTTGATTGTGTGGAACAATTGGATTCTTGTTAATTGAGAGGGTTTGG  1020
Db   872  -------------------------------------------------------------  871

Qy  1021  TTGTTTTTTCAGATCGATGCATTGTGTAAGATAGTCCCGAGACAGGTGAAAGCTGAAAA  1080
                  |||||||||||||||||||||||||||||||||||
Db   872  --------AGATCGATTGCATTGTGTAAGATAGTCCCGAGACAGGTGAAAGCTGAAAA  920

Qy  1081  GGCTGTAACTTGATAAGGGAAAACTGTTGAAGACCTTATTGCTAAGTGTAAAGAAATTGT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   921  GGCTGTAACTTGATAAGGGAAAACTGTTGAAGACCTTATTGCTAAGTGTAAAGAAATTGT  980

Qy  1141  CGAAAGAGAGGCGAAAGAATCAATGATGAGGAGTATGTAAATGATGCTGACCCAAGTAT  1200

---

Db   981  CGAAAGAGAGGCGAAAGAATCAATGATGAGGAGTATGTAAATGATGCTGACCCAAGTAT  1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  1201  CCTGCGTTCTGCTTGCAAGCAGCAGAAGAGAGTTAAACTTTTTCCTTAAGTTTATAAG  1260
           ||||||||||||||||||||||||||||||||
Db  1041  CCTGCGTTCTGCTTGCAAGCAGCAGAAG-------------------------------  1069

Qy  1261  CAAATTGGCCTTCATTATCGCATAATCGAAGCTGATGTGTGCATTGTGAGGGTTTTCAG  1320
Db  1070  ------------------------------------------------------AG  1071

Qy  1321  GTATCAAGTGTGCAGTTACGGGATGATCTTCTCCAATGCTCGTAGCGGGTCATGAAACC  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1072  GTATCAAGTGTGCAGTTACGGGATGATCTTCTCCAATGCTCGTAGCGGGTCATGAAACC  1131

Qy  1381  ACTGGATCGTCCTCACTTGACACTTATCTCCTAAGTAAGTAACTTAATGTATCTTC  1440
           ||||||||||||||||||||||||||||||||||||||||
Db  1132  ACTGGATCGTCCTCACTTGACACTTATCTCCTAAGTA-------------------  1171

Qy  1441  TACTTTGCTATGCTAGAGAATTTACTTGGATGGGACGCTTCTCGTTCTCATTTACCTCTT  1500
Db  1172  -----------------------------------------------------------  1171

Qy  1501  CAAATTCTCTATGTTCATAGAACTCATCTGCATTAAGGAAAGCACAGAAGAAGTAGACA  1560
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1172  --------AGAACTCATCTGCATTAAGGAAAGCACAGAAGAAGTAGACA  1213

Qy  1561  GAGTGTTAGAAGGAAGAAAACCCGGCTTCGAGGATATAAAGGAGTTGAAGTACATCACTC  1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1214  GAGTGTTAGAAGGAAGAAAACCCGGCTTCGAGGATATAAAGGAGTTGAAGTACATCACTC  1273

Qy  1621  GTTGTATAAACGAGTCAATGCGTCTCATCCTCGTGAAGCAATCAAGCTCATC  1680
           |||||||||||||||||||||||||||||||||||||
Db  1274  GTTGTATAAACGAGTCAATGCGTCTCATCCTCATCCTCCT-----------  1314

Qy  1681  TCTCTAATTATTCATGAACTAAATTTTCTGATTGATTTGTTTCCTGTAGTCTTGATAA  1740
                                                              |||||||||
Db  1315  --------------------------------------------GTCTTGATAA  1324

Qy  1741  GAAGAGCTCAAGTTCCTGACATTCTTCCTGGAACTATAAGTCAATACCGGACAAGACA  1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1325  GAAGAGCTCAAGTTCCTGACATTCTTCCTGGAACTATAAGTCAATACCGGACAAGACA  1384

Qy  1801  TTATGATTTCAGTCTATAACATCATCGTTCTTCCGAGGTA  1841
           ||||||||||||||||||||||||||||||||||||||||
Db  1385  TTATGATTTCAGTCTATAACATCATCGTTCTTCCGAGGTA  1425

```
            GenCore version 5.1.6
      Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   January 12, 2006, 20:19:34 ; Search time 12340 Seconds
                                         (without alignments)
                                         11364.084 Million cell updates/sec

Title:          US-10-751-235-5
Perfect score:  2467
Sequence:       1 atggagtcctcactctttc.........tgaaggtgagccaaggtaa 2467

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       5883141 seqs, 2842172653 residues

Total number of hits satisfying chosen parameters:   11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
             1: gb_ba:*
             2: gb_in:*
             3: gb_env:*
             4: gb_om:*
             5: gb_ov:*
             6: gb_pat:*
             7: gb_ph:*
             8: gb_pr:*
             9: gb_ro:*
            10: gb_sts:*
            11: gb_sy:*
            12: gb_un:*
            13: gb_vi:*
            14: gb_htg:*
            15: gb_pl:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | | Score | Query Match % | Length | DB | ID | Accession | Description |
|---|---|---|---|---|---|---|---|---|
| 1 | | 2467 | 100.0 | 92611 | 15 | ATT4D2 | AL132958 | Arabidops |
| 2 | | 803 | 32.5 | 1620 | 15 | AY424805 | AY424805 | Arabidops |
| 3 | | 803 | 32.5 | 1858 | 15 | AY091083 | AY091083 | Arabidops |
| 4 | | 803 | 32.5 | 3304 | 15 | AF367289 | AF367289 | Arabidops |
| 5 | | 376.4 | 15.3 | 378 | 15 | ATH526463 | AJ526463 | Arabidops |
| 6 | | 268.6 | 10.9 | 1077 | 15 | AK220829 | AK220829 | Arabidops |
| 7 | | 252 | 10.2 | 1967 | 15 | BT012891 | BT012891 | Lycopersi |
| 8 | | 239.6 | 9.7 | 110000 | 15 | AP008216_208 | AP008216_208 | Continuation (209 |
| 9 | | 239.6 | 9.7 | 178022 | 15 | AC025783 | AC025783 | Oryza sat |
| 10 | | 239.6 | 9.7 | 299735 | 15 | AE017117 | AE017117 | Oryza sat |
| 11 | | 183 | 7.4 | 1876 | 15 | AK065689 | AK065689 | Oryza sat |
| 12 | | 146.8 | 6.0 | 205408 | 14 | AC155564 | AC155564 | Zea mays |
| 13 | | 91 | 3.7 | 1788 | 15 | AY142017 | AY142017 | Arabidops |
| 14 | | 91 | 3.7 | 2017 | 15 | AY058173 | AY058173 | Arabidops |
| 15 | | 91 | 3.7 | 2057 | 15 | AY056446 | AY056446 | Arabidops |
| 16 | | 83.4 | 3.4 | 4217 | 15 | AK068163 | AK068163 | Oryza sat |
| 17 | | 75.2 | 3.0 | 178657 | 14 | AC155503 | AC155503 | Zea mays |
| 18 | | 74.8 | 3.0 | 110000 | 14 | CR954209_3 | CR954209_3 | Continuation (4 of |
| 19 | | 73.2 | 3.0 | 1831 | 15 | AF022457 | AF022457 | Glycine m |
| 20 | | 72.8 | 3.0 | 1926 | 15 | PSCYTP450 | 249263 | P.sativum m |
| 21 | | 72 | 2.9 | 1886 | 15 | BT002582 | BT002582 | Arabidops |
| 22 | | 72 | 2.9 | 1921 | 15 | AY062675 | AY062675 | Arabidops |
| 23 | | 71.8 | 2.9 | 110000 | 14 | CR954213_4 | | Continuation (5 of |
| 24 | c | 71.4 | 2.9 | 110000 | 14 | CR954201_08 | | Continuation (9 of |
| 25 | | 71.2 | 2.9 | 625 | 15 | AK221212 | AK221212 | Arabidops |
| 26 | c | 70.4 | 2.9 | 119420 | 15 | AC079041 | AC079041 | Arabidops |
| 27 | | 67 | 2.7 | 2093 | 15 | AK100596 | AK100596 | Arabidops |
| 28 | | 67 | 2.7 | 2401 | 15 | AK066680 | AK066680 | Oryza sat |
| 29 | | 66.4 | 2.7 | 181167 | 14 | AC155486 | AC155486 | Zea mays |
| 30 | | 66.2 | 2.7 | 2038 | 15 | AY601887 | AY601887 | Ginkgo bi |
| 31 | c | 65.2 | 2.6 | 103973 | 15 | AP004028 | AP004028 | Oryza sat |
| 32 | c | 65.2 | 2.6 | 110000 | 15 | AP008208_351 | | Continuation (352 |
| 33 | c | 65.2 | 2.6 | 145014 | 15 | AP004048 | AP004048 | Oryza sat |
| 34 | c | 59.4 | 2.4 | 197405 | 15 | AL161540 | ATCHRIV40 | Arabidops |
| 35 | c | 59.4 | 2.4 | 202860 | 15 | Z97337 | ATFCA2 | Arabidops |
| 36 | | 59 | 2.4 | 109093 | 15 | AC148994 | AC148994 | Medicago |
| 37 | c | 59 | 2.4 | 120761 | 15 | AC124218 | AC124218 | Medicago |
| 38 | | 58.2 | 2.4 | 110000 | 14 | CR954215_1 | | Continuation (2 of |
| 39 | | 57.8 | 2.3 | 404 | 6 | BD224432 | BD224432 | Materials |
| 40 | | 57.8 | 2.3 | 404 | 6 | AR216482 | AR216482 | Sequence |
| 41 | | 57.8 | 2.3 | 404 | 6 | AR432884 | AR432884 | Sequence |
| 42 | | 56.6 | 2.3 | 2269 | 15 | AF459441 | AF459441 | Skeletone |
| 43 | c | 55.2 | 2.2 | 207718 | 14 | AC155487 | AC155487 | Zea mays |
| 44 | c | 53.8 | 2.2 | 86539 | 15 | AP004485 | AP004485 | Lotus cor |
| 45 | c | 53 | 2.1 | 110000 | 1 | CP000001_30 | | Continuation (31 o |

ALIGNMENTS

```
RESULT 1
ATT4D2
LOCUS       ATT4D2          92611 bp    DNA     linear   PLN 16-APR-2005
DEFINITION  Arabidopsis thaliana DNA chromosome 3, BAC clone T4D2.
ACCESSION   AL132958
VERSION     AL132958.1  GI:643245
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Nyakatura,G., Fartmann,B., Dauner,D., Sterr,W., Holland,R.,
            Weichselgartner,M., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F.
            and Salanoubat,M.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 92611)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de mayer@mips.biochem.mpg.de Project
            Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
            d'interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue
            Gaston Cremieux, BP191, 91006 Evry Cedex, France;
            http://www.genoscope.cns.fr
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1..92611
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /chromosome="3"
                     /ecotype="Columbia"
     gene            2081..2646
                     /gene="T4D2.10"
     CDS             join(2081..2434,2533..2646)
                     /gene="T4D2.10"
```

```
                     /note="similarity to BETA-GALACTOSIDASE PRECURSOR,
                     Lycopersicon esculentum, gb:P48980"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB64211.1"
                     /db_xref="GI:6630728"
                     /db_xref="GOA:Q9SCQ3"
                     /db_xref="InterPro:IPR000922"
                     /db_xref="UniProt/TrEMBL:Q9SCQ3"
                     /translation="METLQCRHQHVFILLLVLFHSSLFGLASKIDVSDDARGIRIDGG
                     QKRFLTNSPQHGKEHAACTNEEPDLGPLTRISCNEPGYVITKINFADYGNPTGTCCHF
                     RRGNCGARATMRIVKKNCLGKEKCHLLVTDEMFGPSKCKGAPMLAVBTTCTIA"
exon             2081. .2434
                     /gene="T4D2.10"
                     /number=1
intron           2435. .2532
                     /gene="T4D2.10"
                     /number=1
exon             2533. .2646
                     /gene="T4D2.10"
                     /number=2
gene             3341. .8521
                     /gene="T4D2.20"
CDS              join(3341. .3364,3812. .3934,4065. .4305,4394. .4956,
                     5131. .6090,6168. .6490,6606. .6884,6997. .7249,7359. .7529,
                     7645. .7737,7816. .7923,8048. .8164,8257. .8346,8438. .8521)
                     /gene="T4D2.20"
                     /note="similarity to KIAA0010 protein, Homo sapiens,
                     gb:A38919"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB64212.1"
                     /db_xref="GI:6630729"
                     /db_xref="GOA:Q9SCQ2"
                     /db_xref="InterPro:IPR000048"
                     /db_xref="InterPro:IPR000569"
                     /db_xref="UniProt/TrEMBL:Q9SCQ2"
                     /translation="MDLNRKHKVSLRGASSGEISRDALLAKVSQERELRSYARRANAA
                     SLFIQRVWRSYIVRKKAAIEIQEEWENLLSCHSVTLTKSWVSSRVLRPFLFFVRSLSV
                     QHQKIQAREIHCMQTCFKILLESINSNDQGYNFCSLAVGTSEDSKTWACQTRRMVSLC
                     SFLLTECNYSQERIKDVIGVNALLLRILIVLTDPKSWKIITNENFEDAETAKKIIIQF
                     IGSCKSGYYSAVRRYIKTLTKGHTDERLVITTSAVTLALRPFHVKQPAFVDDNQPDTNL
                     AVEEYVSLILTIPRLVCYLPSALIRALKHKSILMPSFHTILLLKDKILNIISEMENSB
                     KQSCTMEIPSVGWVIGNIISLATVSETDFMDPQESNPEMFYVLYVHVIVTLAENLLSQ
                     VESVGIQDIHLDIEATSNETEKGNSVKISFVEMLRPVCQQWHLAKLLAASGKEIRVIA
                     DKDASTSSKKGSETLGLLDIARLYSCMLRIFCVMNPVLGPLPVLNMLSFCPGYIVSLW
                     NSLESVLLPENGCTADDASHGSAKTSWNTRSPSEKKLCKHLKNDSVNKWVNVLNKFSGK
                     SPGPREHVECTSDQPSGQVNESTNDVWDVETLRGGPVGISKEVSCLLHLFCATYAHL
                     LVVLDDIQFYBKQVPFMLEKQQRIASMLNTLVYYGLLRGTGPESRQLMDSAIRCLHLL
                     YERDCRHPFCASALWLSPGRTSRPPIAPAARTHEVLPTSDVLTTPSMGSVITITPHVF
                     PFEERVHVFREFISKDKASRKMAGEVDAPGARSIEIVVRRGHVVEDGFQQLNSIGSRL
                     KSSIHVSFVNESGLPEAGLDYGGLSKEFLTDITKAAPATEYGLFSQTPTSDRLLVPSP
                     SARHLRNGIQMIEFLGRIVGKALYEGILLDYSFSHVFIQKLLGRYSFIDELSGLDPEL
                     YRNLMYVKHYDGDLKELCLDFTVTERFCGKMSIIELXPGGKDTSVTNENKMQYIHAMA
                     DYKLNRQIVPFSNAFYRGLTDLISPAWLKLFNAHEFNQLLSGGNHDIDVDDLRRNTKY
                     TGGYSDSSRTIKIFWEVMKGFEPSERCLLLKFVTSCSRAPILGFKYIQPTFPIIHKVSC
                     DTSLWAAIGGQDVBRLPSASTCYNTLKLPTYKRASTMRBKLLYAITSNAGFELS"
exon             3341. .3364
                     /gene="T4D2.20"
                     /number=1
intron           3365. .3811
                     /gene="T4D2.20"
                     /number=1
exon             3812. .3934
                     /gene="T4D2.20"
                     /number=2
intron           3935. .4064
                     /gene="T4D2.20"
                     /number=2
exon             4065. .4305
                     /gene="T4D2.20"
                     /number=3
intron           4306. .4393
                     /gene="T4D2.20"
                     /number=3
exon             4394. .4956
                     /gene="T4D2.20"
                     /number=4
intron           4957. .5130
                     /gene="T4D2.20"
                     /number=4
exon             5131. .6090
                     /gene="T4D2.20"
                     /number=5
intron           6091. .6167
                     /gene="T4D2.20"
                     /number=5
exon             6168. .6490
                     /gene="T4D2.20"
                     /number=6
intron           6491. .6605
                     /gene="T4D2.20"
                     /number=6
exon             6606. .6884
                     /gene="T4D2.20"
                     /number=7
intron           6885. .6996
                     /gene="T4D2.20"
                     /number=7
exon             6997. .7249
                     /gene="T4D2.20"
                     /number=8
intron           7250. .7358
                     /gene="T4D2.20"
                     /number=8
exon             7359. .7529
                     /gene="T4D2.20"
                     /number=9
intron           7530. .7644
                     /gene="T4D2.20"
                     /number=9
exon             7645. .7737
                     /gene="T4D2.20"
                     /number=10
exon             7738. .7815
                     /gene="T4D2.20"
                     /number=10
intron           7816. .7923
                     /gene="T4D2.20"
                     /number=11
exon             7924. .8047
                     /gene="T4D2.20"
                     /number=11
intron           8048. .8164
                     /gene="T4D2.20"
                     /number=12
exon             8165. .8256
                     /gene="T4D2.20"
                     /number=12
intron           8257. .8346
                     /gene="T4D2.20"
                     /number=13
exon             8347. .8437
                     /gene="T4D2.20"
                     /number=13
intron           8438. .8521
                     /gene="T4D2.20"
                     /number=14
exon             9057. .10601
                     /gene="T4D2.30"
gene             complement(join(9057. .9151,9403. .9658,9810. .10043,
                     10124. .10254,10370. .10601))
                     /gene="T4D2.30"
gene             complement(join(9057. .9151,9403. .9658,9810. .10043,
                     10124. .10254,10370. .10601))
                     /gene="T4D2.30"
CDS
```

```
              /note="strong similarity to proline-rich protein, Raphanus
              sativa, PIR:S16748"
              /product="putative protein"
              /protein_id="CAB64213.1"
              /db_xref="GI:6630730"
              /db_xref="GOA:Q9SCQ1"
              /db_xref="InterPro:IPR001087"
              /db_xref="UniProt/TrEMBL:Q9SCQ1"
              /translation="MRVSGFPVLLLVSCFPCKSKGAVVPALIMPGDSIVDVGNNNNLL
              SIVKSNFLPYGRDFIDQRPTGRFCNGKLAVDFSAEYLGFSSYPPAFLSRBASNENILI
              GANFASASSGYDATSVPFGISLTRQLSYRAYQNRVTRMIGRGNARILFSRGIHIL
              SAGSSDFLQNYYINPLLNLNTPDQFADILLRSFSEFIQNLYELGARRIGVISLPPMG
              CLPAAITLFGAGNKSCVERLNNDAIMPNTKLENTTRLILMNRHSGLRLVAFNVYQPPLD
              IITNPTDNGYVFWDGPHPTRAVNELLAGQLLGQGISLIN"

Query Match      100.0%;   Score 2467;  DB 15;  Length 92611;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 2467;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy     1 ATGGAGCTTCACTCTTTTCTCCATCTTCCTCTCTTCCTCTTTCTTCTTCTCACTGCAAATGT 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 16880 ATGGAGCTTCACTCTTTTCTCCATCTTCCTCTCTTCCTCTTTCTTCTTCTCACTGCAAAA 16939

Qy    61 CCTACGCGTCTTTATCACCAAAACCCAAATTCACATTCTCCATCAGATCCTCCATTGAG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 16940 CCTACGCGTCTTTATCACCAAAACCCAAATTCACATTCTCCATCAGATCCTCCATTGAG 16999

Qy   121 AAACCCAAACCCAAACTCGAGACCAATCGAAATCATCGAAATCCAATCATGGTCAGTCCCGAT 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17000 AAACCCAAACCCAAACTCGAGACCAATCGAAATCATCGAAATCCAATCATGGGTCAGTCCCGAT 17059

Qy   181 TGGCTCACAACACTCACTCGTACCCTTCTCCAGGAAAAAACGACGAGTCAGGTATACCA 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17060 TGGCTCACAACACTCACTCGTACCCTTCTCCAGGAAAAAACGACGAGTCAGGTATACCA 17119

Qy   241 ATCGGCGAACGGAAGCTGACGATGTCGACGATGTCGCTCCTCGGAGGTGCTCTCTTCTACCT 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17120 ATCGGCGAACGGAAGCTGACGATGTCGACGATGTCGCTCCTCGGAGGTGCTCTCTTCTACCT 17179

Qy   301 CTCTACAAATGGATGAATGAGTACGGACCCCATTACCGTCGTCGTGGTCTCGTCGTAAT 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17180 CTCTACAAATGGATGAATGAGTACGGACCCCATTACCGTCGTCGTGGTCTCGTCGTAAT 17239

Qy   361 TTCGTAATTGTGAGCGACCCAGCGATAGTAAACATGTTTGAGGAATTATCCAAAGTAC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17240 TTCGTAATTGTGAGCGACCCAGCGATAGTAAACATGTTTGAGGAATTATCCAAAGTAC 17299

Qy   421 GCTAAAGGCTTAGTGCTGAAGTCTCGAAGTTCTATTGGTTCGGGTTTCGGGTTTCGCATCGCT 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17300 GCTAAAGGCTTAGTGCTGAAGTCTCGAAGTTCTATTGGTTCGGGTTTCGGGTTTCGCATCGCT 17359

Qy   481 GAAGGACCTCTTTGGACAGTAAATTCATCCTCCTATCTCAATTTGAAGTTTTTTGGAA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17360 GAAGGACCTCTTTGGACAGTAAATTCATCCTCCTATCTCAATTTGAAGTTTTTTGGAA 17419

Qy   541 TTGTTGGAAGTAATGTGTGACTGTCTCTGTATGATAAGTAACCTCTAATTTTAGGGTTTTAGAT 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17420 TTGTTGGAAGTAATGTGTGACTGTCTCTGTATGATAAGTAACCTCTAATTTTAGGGTTTTAGAT 17479

Qy   601 TCCAATCTTCTCTATTGGGCTTAGCTGAAGTCTGAGTCTGATTTTTTACATAGGCGGAGG 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17480 TCCAATCTTCTCTATTGGGCTTAGCTGAAGTCTGAGTCTGATTTTTTACATAGGCGGAGG 17539

Qy   661 CGGTGGTTCCATCGCTTCACAGGAGGTATTGTCTGTGATTGTGTGAGAGAGTATTCTCGA 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17540 CGGTGGTTCCATCGCTTCACAGGAGGTATTGTCTGTGATTGTGTGAGAGAGTATTCTCGA 17599

Qy   721 AATGTGCAGAGAGAGCCTTGTTGAGAAGTTGCAGCCTTATGCAGAAGACCGGAAGTGCTGTGA 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17600 AATGTGCAGAGAGAGCCTTGTTGAGAAGTTGCAGCCTTATGCAGAAGACCGGAAGTGCTGTGA 17659

Qy   781 ATATGGAAGCGAAGTTCTCTCAGATGACACTGATGGTGTCATTGGGTTGTGTCTCTTTTAACT 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17660 ATATGGAAGCGAAGTTCTCTCAGATGACACTGATGGTGTCATTGGGTTGTGTCTCTTTTAACT 17719

Qy   841 ACAATTCGATTCTTGACTACTGACTAGTCCGTCATTGAAGCTGTTACACTGCTCTTA 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17720 ACAATTCGATTCTTGACTACTGACTAGTCCGTCATTGAAGCTGTTACACTGCTCTTA 17779

Qy   901 AAGAAGCTGAGCTTCGTTCGTTCTCTGCCATATTGGAAGGCAAGTTCCTGTGT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17780 AAGAAGCTGAGCTTCGTTCGTTCTCTGCCATATTGGAAGGCAAGTTCCTGTGT 17839

Qy   961 TTTTCTGTTGGTTGTTGATTGTGTGGAACAATTGGATTCTTGTTAATTGAGAGGGTTTGG 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17840 TTTTCTGTTGGTTGTTGATTGTGTGGAACAATTGGATTCTTGTTAATTGAGAGGGTTTGG 17899

Qy  1021 TTGTTTTTTCAGATGCGATGCATTGTAAGATAGTCCCGAGACAGGTGAAAGCTGAAAA 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17900 TTGTTTTTTCAGATGCGATGCATTGTAAGATAGTCCCGAGACAGGTGAAAGCTGAAAA 17959

Qy  1081 GGCTGTAACTTGATAAGGGAAAACTGTTGAAGACCTTATTGCTAAGTGTAAAGAAATTGT 1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17960 GGCTGTAACTTGATAAGGGAAAACTGTTGAAGACCTTATTGCTAAGTGTAAAGAAATTGT 18019

Qy  1141 CGAAAGAAGGCGAAAGAATCAATGATGAGGAGTATGTAAATGCTGACCCAAGTAT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18020 CGAAAGAAGGCGAAAGAATCAATGATGAGGAGTATGTAAATGCTGACCCAAGTAT 18079

Qy  1201 CCTGCGGTTCTTGCTCTGCAAGCAGAGAAGAGAGGTTTAAACTTTTTTCCTTAAGTTTATAAG 1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18080 CCTGCGGTTCTTGCTCTGCAAGCAGAGAAGAGAGGTTTAAACTTTTTTCCTTAAGTTTATAAG 18139

Qy  1261 CAAATTTGGCCTTTCATTATGCGCATAATCGACGAGCTGATGTTGCATTGTGAGGGTTTCAG 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18140 CAAATTTGGCCTTTCATTATGCGCATAATCGACGAGCTGATGTTGCATTGTGAGGGTTTCAG 18199

Qy  1321 GTATCAAGTGTGCAGTTACGGGATGATCTCGTAGCGGTCATGAAACC 1380
           |||||||||||||||||||||||||||||||||||||||||||||||
Db 18200 GTATCAAGTGTGCAGTTACGGGATGATCTCGTAGCGGTCATGAAACC 18259

Qy  1381 ACTGGATCGTGTCCTCACTTGACACTTTATCTCCTAGGTAAGTACCTTAATGTATCTTC 1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18260 ACTGGATCGTGTCCTCACTTGACACTTTATCTCCTAGGTAAGTACCTTAATGTATCTTC 18319

Qy  1441 TACTTTGCTATGCTAGAGAATTTACTTGGATGGAATTATCCATTTCTCATTTACCTCTT 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18320 TACTTTGCTATGCTAGAGAATTTACTTGGATGGAATTATCCATTTCTCATTTACCTCTT 18379

Qy  1501 CAAATTCTCTATGTTCATAGAACTCATCGTAAATTTCTGATTGGTTCCTCGGTTCGATAA 1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18380 CAAATTCTCTATGTTCATAGAACTCATCGTAAATTTCTGATTGGTTCCTCGGTTCGATAA 18439

Qy  1561 GAGTGTTAGAAGGAAGAAACCCGGCTTTCGAGGATATAAAGGAGTTGAAGTACATCACTC 1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18440 GAGTGTTAGAAGGAAGAAACCCGGCTTTCGAGGATATAAAGGAGTTGAAGTACATCACTC 18499

Qy  1621 GTTGTATAAACGAGTCAATGCGTCTCATCCTCCTGTAAGCAATCAAGCTCATC 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18500 GTTGTATAAACGAGTCAATGCGTCTCATCCTCCTGTAAGCAATCAAGCTCATC 18559

Qy  1681 TCTCTAATTATTCATGAACTAAATTTCTGATTGTTTCCTCGGTTCTTGATAA 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18560 TCTCTAATTATTCATGAACTAAATTTCTGATTGTTTCCTCGGTTCTTGATAA 18619

Qy  1801 TTATGAATTTCAGTCTATAACATCCATAGTTCTTCCGAGGTACAGTTCTCTTCTTCCTTCTC 1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 18680 TTATGAATTTCAGTCTATAACATCCATAGTTCTTCCGAGGTACAGTTCTCTTCTTCCTTCTC 18739

Qy  1861 GTCCATAGTAATAACATAGGGGAGCCTAATCCTTCTTCAATGATCTTTGTGTTCGG 1920
```

```
Qy   721 AATGTGCAGAGAGAGCCTTGTTGAGAAGTTGCAGCCTTATGCAGAAGACCGGAAGTGCTGTGA 780
Db 17600 AATGTGCAGAGAGAGCCTTGTTGAGAAGTTGCAGCCTTATGCAGAAGACCGGAAGTGCTGTGA 17659

Qy  1741 GAAGAGCTCAAGTTCCTGACATTCTTCGACATTCTTCCTGGGAACTATAAGGTCAATACCGACAAGACA 1800
Db 18620 GAAGAGCTCAAGTTCCTGACATTCTTCGACATTCTTCCTGGGAACTATAAGGTCAATACCGACAAGACA 18679
```

```
Db   18740 GTCCATAGTATAACATAGGGAGCTAATCCTTCTTCTTCAATGATCTTTGTGTGGTTCGG 18799
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1921 ATATCTAACCGAGTGGACATTCCTAGTATTACATTCATGCCCACATTCTTATGTGTTT 1980
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   18800 ATATCTAACCGAGTGGACATTCCTAGTATTACATTCATGCCCACATTCTTATGTGTTT 18859

Qy    1981 GTTGTTTGTTATTCCAAAGGTATGGGAAAAAGCTGAGGAATTTCTGCCTGAACGATTCGA 2040
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   18860 GTTGTTTGTTATTCCAAAGGTATGGGAAAAAGCTGAGGAATTTCTGCCTGAACGATTCGA 18919

Qy    2041 CATAGATGGCGCAATCCCTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGACAC 2100
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   18920 CATAGATGGCGCAATCCCTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC 18979

Qy    2101 ATCTTTTGACACAAACTACTGAATCAAGATTAGTGGTTTGATTAGGGAATTTAAAAGAT 2160
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   18980 ATCTTTTGACACAAACTACTGAATCAAGATTAGTGGTTTGATTAGGGAATTTAAAAGAT 19039

Qy    2161 GATTTCTTTTTTCACCAGATTCATCCCATTCAGTCGAGGGCCTAGAAAATGTGTAGGCG 2220
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   19040 GATTTCTTTTTTCACCAGATTCATCCCATTCAGTCGAGGGCCTAGAAAATGTGTAGGCG 19099

Qy    2221 ATCAGTTTGCATTGATGGAGGCAATTGTGGCACTCGCGGTGTTTCTTCAGCGGTTAAACG 2280
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   19100 ATCAGTTTGCATTGATGGAGGCAATTGTGGCACTCGCGGTGTTTCTTCAGCGGTTAAACG 19159

Qy    2281 TTGAGCGTGGTTCCTGATCAGACCATTAGCATGAGCAGAGCAACCATACACACCACCA 2340
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   19160 TTGAGCGTGGTTCCTGATCAGACCATTAGCATGAGCAGAGCAACCATACACACCACCA 19219

Qy    2341 ATGTATGCCAATGTTCTCACACTCGAGAGATTAATGAGAGTGTCTGTTTGTTTAGAATG 2400
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   19220 ATGTATGCCAATGTTCTCACACTCGAGAGATTAATGAGAGTGTCTGTTTGTTTAGAATG 19279

Qy    2401 ATTCCAATTTTCTTAAATGCTGATATTTCAATTTCAGGGATTGTATGAAGGTGAGCCA 2460
                   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   19280 ATTCCAATTTTCTTAAATGCTGATATTTCAATTTCAGGGATTGTATGAAGGTGAGCCA 19339

Qy    2461 AAGGTAA 2467
                   |||||||
Db   19340 AAGGTAA 19346
```

```
RESULT 2
AY424805
LOCUS        AY424805              1620 bp    mRNA    linear    PLN 07-JAN-2004
DEFINITION   Arabidopsis thaliana chloroplast carotenoid epsilon-ring
             hydroxylase (LUT1) mRNA, complete cds; nuclear gene for chloroplast
             product.
ACCESSION    AY424805
VERSION      AY424805.1   GI:40218378
KEYWORDS     .
SOURCE       Arabidopsis thaliana (thale cress)
  ORGANISM   Arabidopsis thaliana
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
             rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1  (bases 1 to 1620)
  AUTHORS    Tian,L., Musetti,V., Kim,J., Magallanes-Lundback,M. and
             DellaPenna,D.
  TITLE      The Arabidopsis LUT1 locus encodes a member of the cytochrome P450
             family that is required for carotenoid epsilon-ring hydroxylation
             activity
  JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 101 (1), 402-407 (2004)
   PUBMED    14709673
REFERENCE    2  (bases 1 to 1620)
  AUTHORS    Tian,L., Musetti,V., Kim,J., Magallanes-Lundback,M. and
             DellaPenna,D.
  TITLE      Direct Submission
  JOURNAL    Submitted (30-SEP-2003) Biochemistry and Molecular Biology,
             Michigan State University, East Lansing, MI 48824, USA
FEATURES             Location/Qualifiers
     source          1..1620
```

```
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="3"
                     /ecotype="Columbia"
     gene            1..1620
                     /gene="LUT1"
     CDS             1..1620
                     /gene="LUT1"
                     /codon_start=1
                     /product="chloroplast carotenoid epsilon-ring hydroxylase"
                     /protein_id="AAR83120.1"
                     /db_xref="GI:40218379"
                     /translation="MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIRSSIEKPKP
                     KLETNSSKSQSWVSPDWLTILRTLSSGKNDESGIPIANAKLDDVADLLGGALFLPLY
                     KMWNEYGPIYRLAAGPRNFVVSDPAIAKHVLRNYPKYAKGLVABVSEPLFGSGFAIA
                     EGPLWTARRRAVVPSLHRRYLSVIVERVFCKCAERLVEKLQPYAEDGSAVNMEAKFSQ
                     MTLDVIGLSLFNYNFDSLTTDSPVIEAVYTALKEAELRSTDLLPYWKIDALCKIVPRQ
                     VKAEKAVTLIRETVEDLIAKCKEIVEREGERINDEYVNDADPSILRFLLASREVSS
                     VQLRDDLLSMLVAGHETTGSVLTWTLYLLSKNSSALRKAQBEVDRVLEGRNPAFEDIK
                     ELKYITRCINESMRLYPHPPVLIRRAQVPDILPGNYKVNTGQDIMISVYNIHRSSEVW
                     EKABEFLPERPDIDGAIPNEINTDFKFIPFSGGPRKCVGDQPALMEAIVALAVFLQRL
                     NVELVPDQTISMTTGATIHTTNGLVMKVSQR"

ORIGIN

Query Match          32.5%;   Score 803;   DB 15;   Length 1620;
Best Local Similarity 73.2%;   Pred. No. 4.7e-182;
Matches 1347;   Conservative     0;   Mismatches   0;   Indels  494;   Gaps   5;
```

```
Qy     1 ATGGAGTCTTCACTCTTTCTCCATCTTCCTCTTCTTCTTCTTCTCTTCACTCGCAAAA 60
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGAGTCTTCACTCTTTCTCCATCTTCCTCTTCTTCTTCTTCTCTTCACTCGCAAAA 60

Qy    61 CCTACGGCGTCTTTATCACCAAAACCCAAATTCACATTCTCCATCGAAATCCTCCATTGAG 120
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CCTACGGCGTCTTTATCACCAAAACCCAAATTCACATTCTCCATCGAAATCCTCCATTGAG 120

Qy   121 AAACCCAAACCAAACTGAGACCAATTCATCGAAATCATGGGTCAGTCCCGAT 180
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AAACCCAAACCAAACTGAGACCAATTCATCGAAATCCCAATCATGGGTCAGTCCCGAT 180

Qy   181 TGGCTCACACACTCACTCGTACCCTTTCCTCAGGAAAAACGACGAGTCAGGTATACCA 240
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGGCTCACACACTCACTCGTACCCTTTCCTCAGGAAAAACGACGAGTCAGGTATACCA 240

Qy   241 ATCGCGAACGCGAAGCTCGACGATGTCGCTGATCCTCCGGAGGTGCTCTCTTCTTACCT 300
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ATCGCGAACGCGAAGCTCGACGATGTCGCTGATCCTCCTCGGAGGTGCTCTCTTCTTACCT 300

Qy   301 CTCTACAAATGGATGAATGAGTACGGACACCCATTACCGTCTCGCTCGGTCCTCGTAAT 360
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CTCTACAAATGGATGAATGAGTACGGACACCCATTACCGTCTCGCTCGGTCCTCGTAAT 360

Qy   361 TTCGTAATTGTGAGCGACCAGCCAGCGATAGCTAAACATGTTTGAGGAATTATCCAAAGTAC 420
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TTCGTAATTGTGAGCGACCAGCCAGCGATAGCTAAACATGTTTGAGGAATTATCCAAAGTAC 420

Qy   421 GCTAAAGGCTTAGTCGCTGAAGTCTCTGAATTTCTATTTGGTTCGGGTTCGCTATGCGCT 480
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCTAAAGGCTTAGTCGCTGAAGTCTCTGAATTTCTATTTGGTTCGGGTTCGCTATGCGCT 480

Qy   481 GAAGGACCTCTTGACAGTAATTCATCTCCTCCTATCCAATTTGAAGTTTTGAA 540
                 |||||||||||||||||||||||||||||||
Db   481 GAAGGACCTCTTGGACA----------------------------------------- 498

Qy   541 TTGTGGAAGTAATGTCGTCGACTGTCTTGTATGATAAGTAAACTCAATTTAGGGTTTAGAT 600
                 -------------------------------------------------------- 498

Qy   601 TCCAATCTTCTATTGGGCTTAGCTGAAGTCGTCGATTTTTACATAGGCGAGGGCGTAGAG 660
                                                          |||||||||||||||
Db   499 ----------------------------------------GCGAGGGCGTAGAG 511
```